Primljen / Received: 7.5.2025.
Ispravljen / Corrected: 16.8.2025.
Prihvaćen / Accepted: 1.9.2025.
Dostupno online / Available online: 10.1.2026.

Predviđanje nesigurnih cestovnih dionica pomoću strojnog učenja

Autori:

Mr.sc. **Riste Ristov**, dipl.ing.građ.
Sveučilište sv. Ćirila i Metoda, Sj. Makedonija
Građevinski fakultet
ristov@gf.ukim.edu.mk
Autor za korespondenciju

Prof.dr.sc. **Slobodan Ognjenović**, dipl.ing.građ.
Sveučilište sv. Ćirila i Metoda, Sj. Makedonija
Građevinski fakultet
ognjenovic@gf.ukim.edu.mk

Prof.dr.sc. **Zlatko Zafirovski**, dipl.ing.građ.
Sveučilište sv. Ćirila i Metoda, Sj. Makedonija
Građevinski fakultet
zafirovski@gf.ukim.edu.mk

Izvorni znanstveni rad

Riste Ristov, Slobodan Ognjenović, Zlatko Zafirovski

Predviđanje nesigurnih cestovnih dionica pomoću strojnog učenja

U ovome je radu opisana metodologija strojnog učenja za predviđanje opasnih cestovnih dionica primjenom ponderiranog indeksa nesreća (Wi). U analizu uključena je 161 cestovna dionica u Sjevernoj Makedoniji, ukupne duljine oko 1300 km, pri čemu su razmotrene 23 varijable svrstane u skupine koje se odnose na cestu, promet, okoliš i nesreće. Utjecaj značajki vrednovan je primjenom šest modela, uz podjelu podataka na skup za učenje i skup za testiranje u omjeru 80 : 20. Primjenom ponderiranog SHAP-a izvedeno je jedinstveno rangiranje varijabli, dok je konačni prediktivni model XGBoost temeljen na 15 ulaznih značajki. Potvrđena učinkovitost modela iznosi R^2 = 0,65, a u sklopu operativne prioritizacije postignut je AUROC = 0,69 pri Wi ≥ 10,13, što nadležnim institucijama omogućuje pravodobnu identifikaciju opasnih dionica i odgovarajuće intervencije.

Ključne riječi:
sigurnost u prometu, strojno učenje, predviđanje, SHAP, ponderirani indeks nesreća, analiza prometa

Original research paper

Riste Ristov, Slobodan Ognjenović, Zlatko Zafirovski

Predicting unsafe road sections using machine learning

This paper presents an ML methodology to predict hazardous road segments, using the weighted accident index (Wi). The analysis covers 161 road segments in North Macedonia (~1,300 km) - with 23+1 variables categorized into Road, Traffic, Environmental, and Accident data. Feature influence is evaluated using six models with an 80/20 training/testing split. Weighted SHAP is applied to obtain a single variable ranking; XGBoost with 15 inputs is the final predictor. The model achieves a validated performance (R^2 = 0.65), while operational prioritization yields AUROC = 0.69 at Wi ≥ 10.13, enabling timely identification of hazardous segments and interventions by relevant authorities.

Key words:
road safety, machine learning, prediction, SHAP, weighted accident index, traffic analysis

1. Uvod

Prometne nesreće ozbiljna su globalna prijetnja javnoj sigurnosti, pri čemu posebno ugrožavaju mlađu populaciju. Svjetska zdravstvena organizacija navodi da su prometne nesreće vodeći uzrok smrtnosti među osobama u dobi od 5 do 29 godina, pri čemu godišnje odnosu više od 1,19 milijuna života [1]. Iako su sustavne mjere i infrastrukturna poboljšanja doprinijela smanjenju stopa smrtnosti u razvijenim zemljama, prosječan broj smrtnih slučajeva u Europskoj uniji i dalje je visok i iznosi 45,5 na milijun stanovnika [2].

Stopa smrtnosti u Republici Sjevernoj Makedoniji iznosi 69,5 na milijun stanovnika [3], što je znatno iznad europskog prosjeka i pokazuje zabrinjavajući trend. Ti podaci ukazuju na hitnu potrebu za razvojem novih analitičkih pristupa koji bi unaprijedili sigurnost u prometu, omogućujući pravodobno prepoznavanje visokorizičnih dionica i planiranje odgovarajućih mjera.

Suvremene analitičke tehnike poput strojnog učenja i prostorno-statističkih metoda omogućuju proaktivno otkrivanje nesigurnih cestovnih dionica prije nastanka kritičnih događaja. Primjena tih tehnologija omogućuje identifikaciju opasnih zona na temelju kombiniranog utjecaja više čimbenika, umjesto oslanjanja isključivo na povijesne podatke o nesrećama.

Cilj ovog istraživanja bio je razviti metodologiju za predviđanje nesigurnih cestovnih dionica analizom geometrijskih obilježja, stanja kolnika, intenziteta prometa i vanjskih čimbenika. Podaci primijenjeni u ovome istraživanju obuhvaćaju 161 dionicu primarne cestovne mreže, u ukupnoj duljini od približno 1300 km. Napredni modeli strojnog učenja mogu se primijeniti u identificiranju ključnih čimbenika koji utječu na ponderirani indeks nesreća (engl. *weighted accident index* - Wi) i u razvoju prediktivnih alata za poboljšanje sigurnosti u prometu.

Rezultati ranijih istraživanja pokazuju kako različiti čimbenici, uključujući uzdužni nagib, polumjer zakrivljenosti, stanje kolnika, širinu trakova i ograničenu vidljivost, utječu na rizik od prometnih nesreća [4-6]. Modeli strojnog učenja, osobito algoritmi slučajne šume (engl. Random Forest - RF) i pojačavanja gradijenta (engl. Gradient Boosting - GB) s algoritmima XGBoost i CatBoost, sve se češće primjenjuju u analizi takvih parametara zbog svoje sposobnosti rukovanja složenim i nelinearnim odnosima te identificiranja najutjecajnijih varijabli [7, 8].

Slični pristupi primijenjeni su i u području procjene troškova infrastrukture, gdje su se ansambl-metode strojnog učenja poput RF-a i *boosting* modela pokazale učinkovitima u predviđanju složenih interakcija između višestrukih ulaznih varijabli [9]. "Istraživanja su također istaknula primjenu umjetnih neuronskih mreža (ANN) za predikciju učestalosti sudara, osobito kada su podaci ograničeni ili djelomično dostupni; međutim, interpretabilnost ovih modela i dalje je ograničena" [10]. Nekolicina nedavnih istraživanja pokazala je da SHAP (*SHapley Additive exPlanations*) analiza može poslužiti kao dodatni alat za tumačenje rezultata vrlo preciznih modela, pružajući uvid u relativnu važnost čimbenika kao što su obujam prometa, tip ceste i planinski teren [11].

Osim toga razvijeni su kompozitni indeksi i karte za predviđanje rizika, u kojima su podaci kategorizirani na temelju razine utjecaja i rangiranja umjesto jednostavne binarne klasifikacije [12, 13]. Primjena tih metoda omogućila je bolje usmjeravanje resursa i određivanje prioriteta intervencija.

Međutim, većina postojećih analiza provedena je u zemljama s dobro razvijenim sustavima prikupljanja podataka. Za Republiku Sjevernu Makedoniju i okolnu regiju trenutačno ne postoje modeli koji integriraju lokalna ograničenja i manjak sveobuhvatnih podataka [14]. Cilj ovog istraživanja bio je popuniti navedenu prazninu primjenom nekoliko modela strojnog učenja (XGBoost, RF, GB, CatBoost, LightGBM i višeslojni perceptron – MLP) uz postupno smanjivanje broja ulaznih varijabli na temelju njihova doprinosa ponderiranome indeksu nesreća (Wi). Visokorizične cestovne dionice utvrđene su na temelju reprezentativnih nacionalnih podataka.

2. Pregled literature

Istraživanja koja se bave predviđanjem prometnih nesreća i njihove ozbiljnosti sve češće primjenjuju napredne algoritme strojnog učenja i interpretabilne modele za analizu ključnih utjecajnih čimbenika. U ovome poglavlju sažeta su relevantna istraživanja koja primjenjuju napredne metode za predikciju nesreća temeljene na stvarnim podacima, uključujući informacije o geografskome opsegu, broju slučajeva, primijenjenim modelima i izvedbenim metrikama.

U istraživanju provedenome u Kini, Chen i suradnici [15] primijenili su hibridni model *MSCPO-XGBoost* na skupu podataka od 13.000 slučajeva, pri čemu je postignut koeficijent determinacije (R^2) od 0,918. Analizirali su čimbenike povezane s ozbiljnošću sudara kombinirajući optimiranje i strojno učenje. Iranmanesh i sur. [16] primijenili su XGBoost, stablo odluke (engl. *Decision Tree* - DT) i RF modele na podatke iz 784 slučaja nesreća na ruralnim cestama u jednoj provinciji u Iranu, pri čemu je postignuta najveća vrijednost R^2 od 0,873. Primjenom navedenih modela utvrđeni su dijelovi cesta s povećanim rizikom od prometnih nesreća.

U istraživanju u kojemu su primijenjeni podaci iz Južne Koreje Lee i sur. [17] primijenili su interpretativni pristup s proširenjem podataka na 11.689 zapisa primjenom SHAP-a za prepoznavanje utjecaja povezanih s infrastrukturom i postigli R^2 od 0,842. Mengistu i sur. [18] analizirali su 1037 slučajeva, uključujući podatke o vozačima, cestama i okolišu u Etiopiji, primjenom modela *XGBoost*, pri čemu su postigli R^2 od 0,863. U oba istraživanja transparentnost modela u objašnjenju faktora istaknuta je kao ključna prednost.

Alshehri i sur. [19] primijenili su modele DT i RF na skupu podataka od 3228 nesreća u Saudijskoj Arabiji. Površina ispod krivulje (engl. *area under curve* - AUC) iznosila je 0,78 za predviđanje smrtnosti među dobnim skupinama i vrstama nesreća. AUC je mjera za određivanje učinkovitosti nekog modela. Osim toga model s najboljim rezultatima postigao je preciznost od 0,81 i pouzdanost od 0,75, što pokazuje da ima dobro uravnoteženu sposobnost otkrivanja pozitivnih

Tablica 1. Komparativni sažetak reprezentativnih istraživanja

Istraživanje (ref.)	Regija/Opseg	Zadatak	Modeli	Veličina	Metrika
Chen i sur. [15]	Kina (širom zemlje)	Regresija (ozbiljnost)	MSCPO-XGBoost	13.000	R^2 = 0,918
Iranmanesh i sur. [16]	Iran (ruralne ceste)	Regresija (rizik/dionice)	XGBoost, DT, RF	784	R^2(maks.) = 0,873
Lee i sur. [17]	Južna Koreja (na razini države)	Regresija + XAI	Interpretabilno strojno učenje + SHAP	11.689	R^2 = 0,842
Mengistu i sur. [18]	Etiopija (na regionalnoj razini)	Regresija (ozbiljnost)	XGBoost	1037	R^2 = 0,863
Alshehri i sur. [19]	Saudijska Arabija (više gradova)	Klasifikacija (rizik od nesreće sa smrtnim posljedicama)	DT, RF	3228	AUC ≤ 0,78; preciznost = 0,81; odziv = 0,75
Ahmed i sur. [20]	Novi Zeland (urbane sredine)	Hibridno (regresija + klasifikacija)	XGBoost, LIME	3146	R^2 = 0,839; AUC = 0,87
Alpalhão i sur. [22]	Portugal – Lisabon (urbane sredine)	Hibridno (Velika Britanija)	GB	28.649	RMSE = 0,332

slučajeva. Ahmed i sur. [20] analizirali su 3146 incidenata na Novome Zelandu primjenom objašnjivih modela kao što su XGBoost i metodu LIME (engl. *Local Interpretable Model-Agnostic Explanations*), pri čemu je postignut R^2 od 0,839. Najbolji klasifikacijski model postigao je AUC od 0,87, što ističe njegovu dosljednu učinkovitost u predviđanju razina ozbiljnosti nesreća. Vizualizacija utjecaja čimbenika dodatno ilustrira praktičnu primjenjivost tih modela.

Megnidio-Tchoukouegno i Adedeji [21] upotrijebili su bazu podataka STATS19, koja sadržava 45.000 zapisa iz Ujedinjene Kraljevine. Primijenili su modele GB i RF, pri čemu je najbolji model postigao vrijednost R^2 od 0,881. Alpalhão i sur. [22] analizirali su 28.649 slučajeva iz Lisabona primjenom hibridnoga regresijskog/klasifikacijskog modela GB, s korijenom srednje kvadratne pogreške (RMSE) od 0,332. Ta istraživanja posebno su važna za urbana područja, gdje dostupnost podataka omogućuje složeno modeliranje. U istraživanju koje su proveli Guido i sur. obrađeno je 1349 slučajeva iz regije Cosenze u Italiji primjenom XGBoosta, SVM-a i RF-a. Najviši postignuti R^2 iznosio je 0,896, a modeli su primijenjeni za analizu broja uključenih vozila i obilježja ceste. Pomoću geoprostorne analize i strojnog učenja pokazali su primjenjivost modela u ruralnim okružjima.

Xiao i Duan [24] izradili su okvir dubokog učenja za predviđanje više zadataka primjenjujući ulazne podatke iz 10.563 slučaja, pri čemu je postignuta vrijednost R^2 od 0,894 i srednja apsolutna pogreška (MAE) od 0,243. U svojemu istraživanju proveli su detaljnu analizu SHAP kako bi vizualizirali doprinos svake varijable. U analizi ozbiljnosti sudara kombinirani su interpretabilnost i višefunkcionalnost.

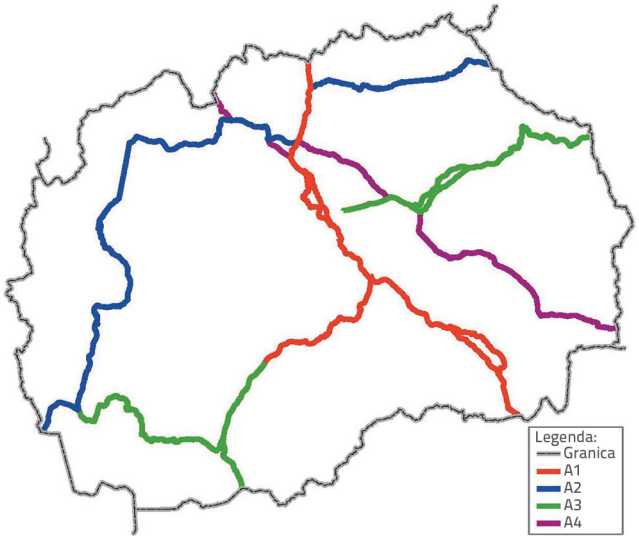
U tablici 1. dan je sažet pregled reprezentativnih istraživanja (regija/opseg, zadatak, modeli, veličina uzorka i metrike) radi lakše i transparentnije usporedbe među istraživanjima i metodološki dosljedne procjene prijavljenih nalaza.

Nadovezujući se na navedena istraživanja, u ovome istraživanju primijenjeno je šest modela strojnog učenja (CatBoost, GB, XGBoost, RF, LightGBM i MLP) treniranih u jednakim uvjetima. Ključni doprinos ovog istraživanja ogleda se u sustavnoj usporedbi učinkovitosti modela u predviđanju ponderiranog indeksa nesreće (Wi) te razvoju metodologije za rangiranje parametara na temelju kombiniranih rezultata važnosti značajki. Osim toga rezultati pružaju praktične smjernice za identifikaciju visokorizičnih dionica unutar ciljnih cestovnih mreža.

3. Pregled podataka

3.1. Opće informacije o cestovnoj mreži

Cestovna mreža Republike Sjeverne Makedonije proteže se na ukupno 14.475 km i obuhvaća autoceste, regionalne i lokalne ceste [25]. Primarna cestovna mreža, duga 897 kilometara, ključna je dionica nacionalne i transeuropske prometne infrastrukture [26]. Uključuje autoceste, brze ceste i dvosmjerne ceste koje osiguravaju glavne prometne veze diljem zemlje i sa susjednim zemljama.



Slika 1. Pregled cesta A-kategorije

Istraživanje je obuhvatilo primarne ceste A1, A2, A3 i A4, koje se razlikuju po svojim tehničkim svojstvima i geometrijskim elementima. Kao što je to prikazano na slici 1., iako službena duljina primarne cestovne mreže iznosi 897 kilometara, analiza obuhvaća približno 1300 kilometara zbog zasebnog razmatranja dvaju cestovnih smjerova s odijeljenim kolnicima (autoceste). Taj pristup omogućuje detaljniju i objektivniju procjenu utjecaja različitih čimbenika na sigurnost cestovnog prometa.

3.2. Opis i obrada podataka

Kombiniranjem vremenske kategorizacije i klasifikacije prema karakteristikama dobiva se sveobuhvatan i sustavan pristup obradi podataka. Vremenska kategorizacija ističe promjene tijekom godina, dok klasifikacija prema karakteristikama omogućuje precizno razumijevanje uloge i utjecaja svakog pojedinog čimbenika. Podaci su obrađeni upotrebom alata Geografskoga informacijskog sustava (GIS), statističkih tehnika i metoda strojnog učenja te su identificirani prostorni i vremenski trendovi.

Konačna analitička baza podataka obuhvaća 161 cestovnu dionicu (≈1300 km) s potpunim zapisima za razdoblje između 2014. i 2023. Prije izrade modela svi su slojevi ponovno projicirani na elipsoid WGS 84, prostorno povezani prema kilometarskim oznakama i provjereni zbog eventualnih dupliciranih identifikatora.

Karakteristike ceste

Ta kategorija uključuje različite geometrijske i funkcionalne parametre kao što su ograničenja brzine, zakrivljenost trase, radijusi krivulja, uzdužni nagibi i nadmorska visina. Analizirana su i bočna opterećenja u zavojima, zaustavni vidni razmak, hrapavost kolnika, dubina kolotraga, koeficijent površinskog trenja te indeks stanja kolnika (PCI). Uključeni su i podaci o gustoći na raskrižju, mostovima i vijaduktima te o stanju vertikalnog i horizontalnog znakovlja [27, 28].

Karakteristike prometa

Intenzitet prometa izražava se godišnjim prosjekom dnevnog prometa (engl. Annual Average Daily Traffic - AADT) na temelju automatskog brojanja prometa s fiksnih i mobilnih uređaja. Pomoću tog parametra moguće je analizirati utjecaj opsega prometa na vjerojatnost nesreća [29].

Karakteristike koje se odnose na okoliš

Klimatski parametri obuhvaćaju prosječne i ekstremne godišnje vrijednosti padalina i temperature, prikupljene tijekom deset godina s odgovarajućih meteoroloških stanica. Podaci su obrađeni geoprostornim metodama kako bi se osigurala pokrivenost visoke rezolucije na razini pojedinačnih cestovnih dionica [30].

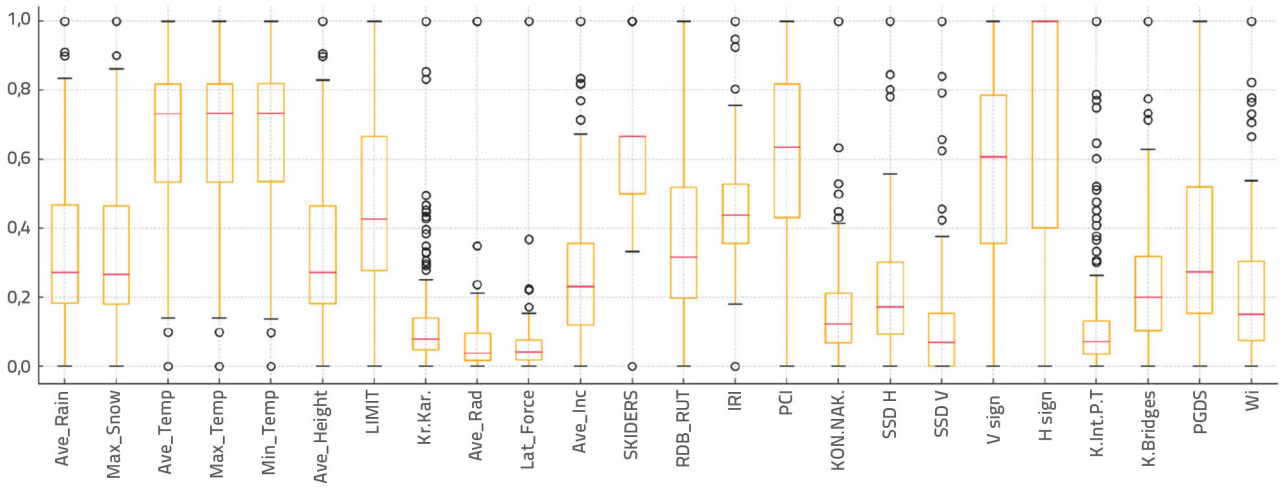
Podaci o prometnim nesrećama

Učestalost i prostorna raspodjela prometnih nesreća analizirani su pomoću indeksa Wi, koji uzima u obzir i broj i ozbiljnost nesreća. Nesreće sa smrtnim posljedicama, nesreće s ozljedama i nesreće samo s materijalnom štetom ponderirane su s vrijednostima 85, 10 i 1, nakon čega je rezultat normaliziran u odnosu na duljinu cestovne dionice. Taj indeks služi kao ključni pokazatelj za usporedbu razina sigurnosti različitih cestovnih dionica [31].

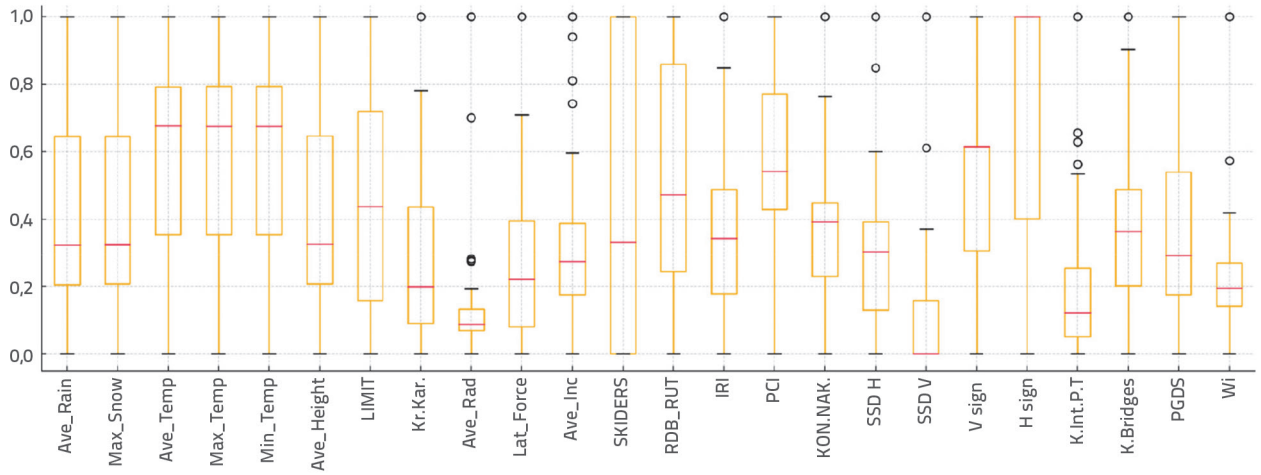
Kontrola kvalitete uključivala je imputaciju manje od 3 % nedostajućih kontinuiranih vrijednosti s medijanom svake varijable, kodiranje kategoričkih pokazatelja jednim korakom i standardizaciju z-vrijednosti svih kontinuiranih ulaza. Geometrijski, prometni i inventarski slojevi preuzeti su sa službenog WebGIS portala *Javnog poduzeća za državne ceste*, čime se osiguravaju dosljednost mjerenja i atribuiranje. Dodatak B pruža potpuni popis definicija varijabli i formula složenih indeksa.

3.3. Statistički sažetak skupa podataka

U sklopu deskriptivne analize primijenjen je podskup za učenje, u kojemu je 80 % podataka normalizirano na raspon od 0 do 1,



Slika 2. Normalizirani kutijasti dijagram ulaznih varijabli i indeksa Wi: skup za učenje



Slika 3. Normalizirani kutijasti dijagram ulaznih varijabli i indeks Wi: testni skup

čime su istaknuti relativni odnosi među ulaznim varijablama i izlaznim parametrima (Wi). Na slici 2. mogu se uočiti izraženi interkvartilni rasponi kod većine varijabli, dok izdvojene točke označavaju cestovne dionice s naglašenim odstupanjima od uobičajenog obrasca, posebno u slučajevima Wi-ja, PGDS-a i PCI-ja.

Preostalih 20 % podataka formiralo je testni podskup, obrađen istim postupkom normalizacije. Drugi dijagram pokazuje da raspodjele zadržavaju oblik i širinu interkvartilnih raspona opaženih u skupu za učenje, što potvrđuje da je podjela podataka reprezentativna i da validacijski postupak ne uvodi sustavna odstupanja u raspodjeli vrijednosti.

Na slici 3. prikazana je distribucija varijabli u testnome podskupu, što omogućuje izravnu usporedbu s podacima za učenje. Dosljednost između skupa za učenje i testnog skupa temelj je za vjerodostojnu procjenu opće upotrebljivosti razvijenih modela. Razlike u duljini kutijastih dijagrama i broju izdvojenih vrijednosti kod pojedinih varijabli upućuju na njihov doprinos varijabilnosti sigurnosnih uvjeta na promatranim cestovnim dionicama.

4. Metodološki pristup za razvoj modela predviđanja ponderiranog indeksa nesreća (Wi)

Za razvoj pouzdanog i interpretabilnog modela za predviđanje ponderiranog indeksa nesreća (Wi) primijenjen je strukturirani postupak koji obuhvaća sedam koraka koji uključuju odabir modela, podešavanje, optimiranje značajki i procjenu. Takav pristup jamči postupan razvoj strukture modela i ulaznih značajki, uz pozorno odvajanje istraživačkog dijela analize od procesa validacije. Tijek rada prikazan je na slici 4.

Početni probir modela

Provedena je početna usporedba devet različitih modela strojnog učenja. To uključuje linearne modele, modele temeljene na DT-u, *boosting* metode, modele temeljene na jezgri (kernel) i neuronske mreže. Ta usporedba omogućila je prepoznavanje algoritama koji pokazuju obećavajući potencijal predviđanja na temelju općih trendova R^2 i metrika pogreške [32, 33].

Odabir modela

Na temelju preliminarnih rezultata modeli koji su imali R^2 veći od 0,50 smatrani su dovoljno pouzdanima za uključivanje u formalni proces validacije.

Podešavanje hiperparametara

Za svaki odabrani model provedeno je 20-iterativno nasumično pretraživanje na 80-postotnome podskupu za treniranje kako bi se identificirali prikladni hiperparametri. U tablici 2. prikazani su hiperparametri nakon podešavanja, primijenjeni u evaluaciji 80/20, pri čemu su sve prikazane metrike izračunane na odvojenome (*held-out*) testnom skupu. U svim je postupcima korištena fiksna vrijednost parametra *random_state* = 42 kako bi se osigurala ponovljivost rezultata.

Odabir značajki i analiza pouzdanosti

Puni skup od 23 ulazna parametra postupno je smanjivan primjenom SHAP-a i metoda temeljenih na permutacijama, pri

Slika 4. Linearan tijek rada za razvoj modela predviđanja ponderiranog indeksa nesreća (Wi)

Tablica 2. Konačni hiperparametri za svaki algoritam (podjela 80/20)

Model	Konačni hiperparametri
XGBoost (konačni prediktor)	n_estimators = 100; max_depth = 4; learning_rate = 0,3 (default); random_state = 42
Pojačavanje gradijenta (GB)	n_estimators = 100; max_depth = 4; learning_rate = 0,1 (default); random_state = 42
Slučajna šuma (RF)	n_estimators = 100; max_depth = 4; min_samples_split = 2; random_state = 42
CatBoost (CB)	iterations = 1000; depth = 6; learning_rate = 0,03; random_state = 42
LightGBM (LGBM)	n_estimators = 300; learning_rate = 0,03; max_depth = 4; min_child_samples = 10; random_state = 42
Višeslojni perceptron (MLP)	hidden_layer_sizes = (100, 50); activation = "relu"; alpha = 0,0005; max_iter = 1000; random_state = 42

čemu je učinkovitost praćena nakon svakog uklanjanja. Konačni odabir uključivao je samo najutjecajnije značajke za svaki model. Stabilnost odabira značajki provjerena je analizama temeljenima na korelaciji različitih *random seed* inicijalizacija i varijanti modela [34, 35]. Važnost varijabli određena je permutacijskim pristupom, pri čemu se bilježilo smanjenje R^2 nakon nasumičnog miješanja svake varijable.

Treniranje i testiranje (podjela 80/20)

S optimiranim hiperparametrima i smanjenim skupom značajki, svaki je model treniran na 80 % podataka i testiran na preostalih 20 %. To je omogućilo nepristrane procjene performansi, temeljene isključivo na odvojenome (*held-out*) testnom skupu [35, 36]. Performanse su prikazane kao R^2 (%), MAE i RMSE za testni skup koji nije primijenjen za treniranje.

Integracija za tumačenje (bez ansambla za predviđanje)

Nije primijenjen zaseban prediktivni ansambl, a integracija je poslužila samo za dobivanje robusnog rangiranja među modelima ponderiranim SHAP-om, pri čemu je doprinos svakog modela proporcionalan njegovoj vrijednosti R^2 [39].

Formulacija konačnih modela

Za konačni model predviđanja odabran je XGBoost s 15 ulaza kao algoritam s najvišim performansama u evaluaciji 80/20, a sve metrike prikazane su na temelju validacije izvan uzorka. Eventualno preoblikovanje modela na cijelome skupu podataka izvodi se isključivo za primjenu na novim cestovnim dionicama, dok svi prijavljeni rezultati ostaju temeljeni na evaluaciji 80/20 [46].

Ta metodologija pruža koherentan i ponovljiv okvir za predviđanje indeksa Wi. Svaki korak, od početnog testiranja modela do konačne implementacije, bio je pozorno strukturiran kako bi se osigurali transparentnost, robusnost i znanstvena strogost. Primjena odabranih modela uz odgovarajuće postupke podešavanja i validacije rezultirala je alatom koji je točan i primjenjiv u praksi za procjenu sigurnosti cestovnog prometa.

5. Analitička procjena učinkovitosti prediktivnih modela

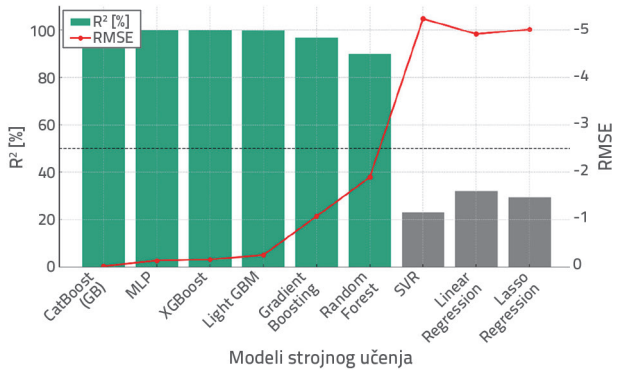
Evaluacija različitih pristupa strojnog učenja za predviđanje ponderiranog indeksa nesreća prikazana je u fazama. Evaluacija obuhvaća R^2 vrijednosti modela tijekom treniranja s punim skupom podataka, testiranje primjenom podjele 80/20, odabir optimalnih parametara i razvoj konačnog prediktivnog modela.

5.1. Indikativna evaluacija treniranja i probir modela

U početnoj fazi analize svi odabrani modeli trenirani su primjenom cijelog skupa podataka. Taj pristup omogućio je brzu procjenu mogućnosti različitih metoda za prediktivno modeliranje ponderiranog indeksa nesreća Wi [43, 45]. Analiza je uključivala različite matematičke pristupe kao što su linearni modeli, modeli temeljeni na stablima odluke (DT), *boosting* modeli, metode temeljene na jezgrenim funkcijama i umjetne neuronske mreže (ANN). Tih devet modela odabrano je na temelju njihove kompatibilnosti s prirodom dostupnih podataka i njihove dokazane primjenjivosti u prethodnim istraživanjima sigurnosti cestovnog prometa [6, 8].

Rezultati treniranja modela odražavaju njihov općeniti potencijal i primijenjeni su za prepoznavanje onih prikladnih za daljnju analizu. Kriterij odabira bilo je postizanje vrijednosti R^2 veće od 50 %, što se smatralo najnižim pragom za obuhvaćanje varijabilnosti u Wi.

Slika 5. ilustrira početne rezultate svih devet modela, prikazane kroz metrike R^2 i RMSE. Modeli s vrijednostima R^2 iznad 50 % smatrani su prikladnima za modeliranje te vrste podataka i zadržani za daljnje potvrđivanje.



Slika 5. Početne vrijednosti R^2 i RMSE za devet prediktivnih modela

Na temelju početne analize odabrano je šest modela, CatBoost [37], MLP, XGBoost [47], LightGBM [38], GB i RF [40], za daljnju evaluaciju primjenom strukturiranog postupka opisanog u

metodologiji. Isključeni modeli nisu zadovoljili prag prediktivne sposobnosti i zato nisu primijenjeni u daljnjim koracima optimiranja.

5.2. Definiranje optimalnog broja utjecajnih parametara

Provedena je kombinirana analiza primjenom vrijednosti SHAP i važnosti permutacije kako bi se razjasnio utjecaj pojedinačnih parametara na Wi. Obje metode omogućuju transparentno tumačenje uloge svake ulazne varijable u konačnome predviđanju, što je ključno za donošenje praktičnih zaključaka i definiranje formula za predviđanje.

SHAP vrijednosti potječu iz teorije igara i izražavaju doprinos svakog parametra određenome predviđanju. Te vrijednosti temelje se na principu pravedne raspodjele utjecaja na rezultat modela. SHAP omogućuje detaljan uvid u to koji parametri imaju najveći utjecaj te je li taj utjecaj pozitivan ili negativan, ovisno o smjeru i veličini vrijednosti [42-44]. Vrijednost parametra *i* može se izračunati na sljedeći način:

$$\Phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\left(|F|-|S|-1\right)!}{|F|!}\left[f\left(S \cup \{i\}\right) - f\left(S\right)\right] \quad (1)$$

gdje Φ_i predstavlja vrijednost SHAP za parametar i, *S* podskup preostalih parametara, a *f*(*S*) rezultat modela za ulazni skup *S*.

Osim toga SHAP metodologija omogućuje neizravno promatranje interakcija između parametara putem njihova kumulativnog učinka na rezultat modela.

Za provjeru tih rezultata provedena je analiza važnosti permutacijama. Taj pristup procjenjuje važnost svakog parametra mjerenjem promjene u učinkovitosti modela kada se vrijednosti određenog parametra zamijene permutiranim vrijednostima. Ako ta zamjena uzrokuje znatno smanjenje preciznosti, tada se parametar smatra vrlo utjecajnim [40].

Utjecaj parametra definiran je razlikom u vrijednostima R^2, izraz (2):

$$\text{Importance}_i = R^2_{\text{original}} - R^2_{\text{permuted,i}} \quad (2)$$

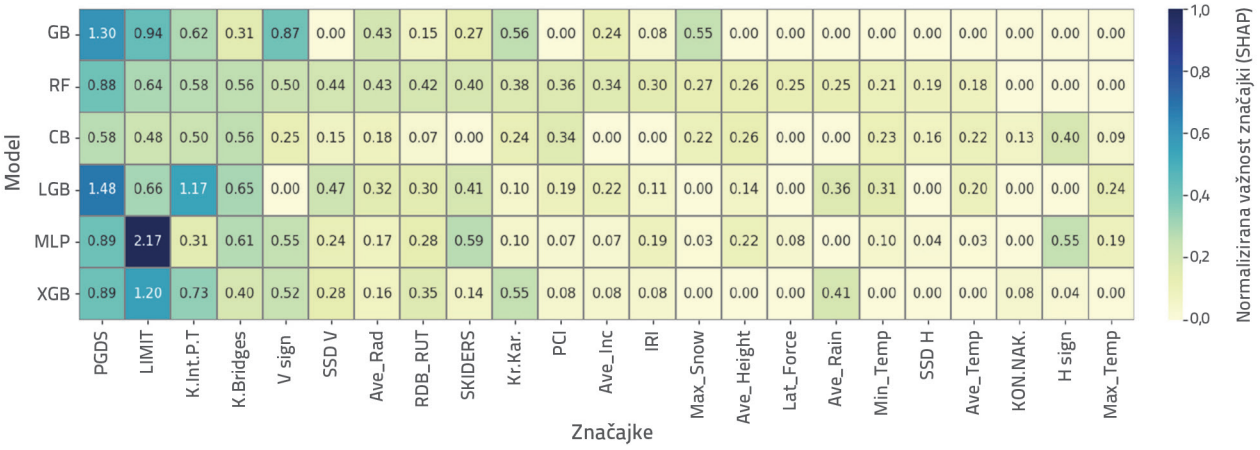
pri čemu je R^2_{original} objašnjena varijanca izvornog modela a $R^2_{\text{permuted,i}}$ vrijednost dobivena nakon permutacije vrijednosti parametra *i*. Što je veća razlika, to je parametar utjecajniji na predviđanja modela.

Za pet od šest modela (GB, RF, CatBoost, LightGBM i XGB) upotrijebljen je isti SHAP objašnjivač temeljen na paketu TreeExplainer, dok je za MLP model primijenjen objašnjivač KernelExplainer zbog svoje strukture neuronske mreže.

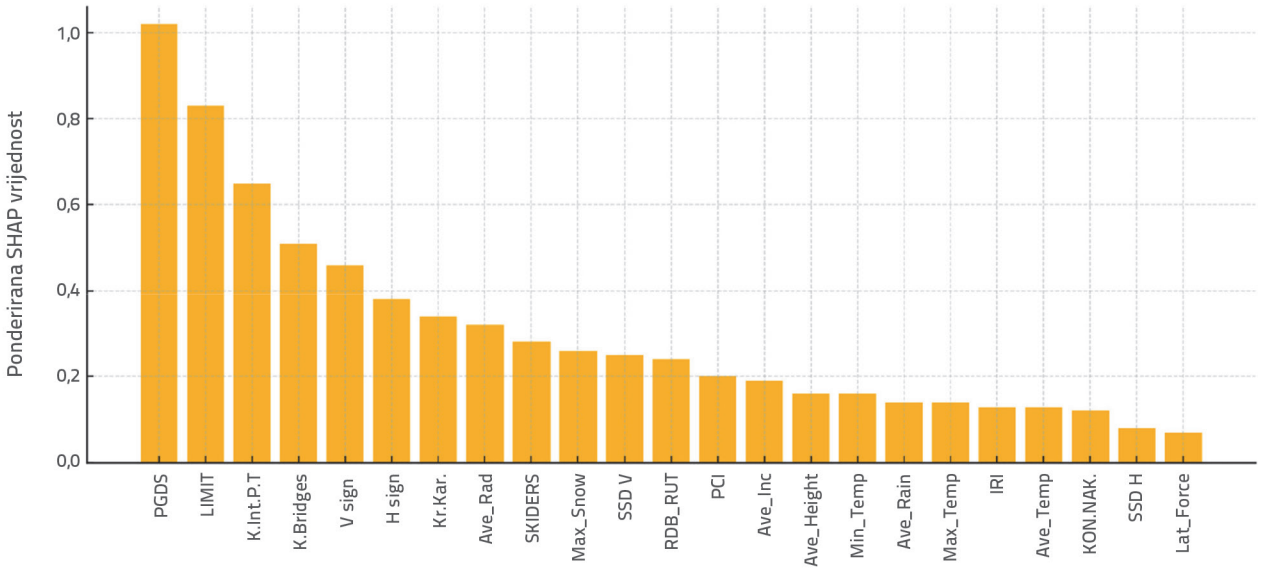
Na slici 6. prikazana je normalizirana važnost svakog od 23 parametra u usporedbi među modelima. Brojevi u kvadratima apsolutne su SHAP vrijednosti zaokružene na dvije decimale, dok boja vizualno označava relativnu važnost na skali od 0 do 1. Ta vizualizacija omogućuje izravnu usporedbu utjecaja parametara u svim modelima; PGDS, LIMIT i K.Int. P.T nedvojbeno predstavljaju najdosljednije utjecajne parametre. Istodobno parametri poput KON.NAK., H sign i Max_Temp pokazuju nižu ili selektivnu važnost samo u pojedinačnim modelima. Za konačno rangiranje parametara SHAP vrijednosti dobivene iz šest modela ponderirane su njihovim pripadajućim vrijednostima R^2 iz analize na skupu podataka podijeljenom u omjeru 80/20. Te su vrijednosti navedene u tablici 3.

Tablica 3. Vrijednosti u pogledu učinkovitosti R^2 primijenjene kao težine u konačnome SHAP izračunu

Model	Vrijednost R^2	Težina
Pojačavanje gradijenta (GB)	0,5381	0,225
Slučajna šuma (RF)	0,5126	0,214
CatBoost (CB)	0,4664	0,195
LightGBM (LGBM)	0,3447	0,144
Višeslojni perceptron (MLP)	0,1607	0,067
XGBoost (XGB)	0,1043	0,044



Slika 6. Normalizirane vrijednosti važnosti značajki prema SHAP metodi za sve modele



Slika 7. Važnost značajki (ponderirane SHAP vrijednosti)

Nadalje, kako bi se rezultati svih modela integrirali u sintetičku procjenu, za svaki parametar izračunana je ponderirana SHAP vrijednost na sljedeći način:

$$S_j = w_1 \cdot S_{j1} + w_2 \cdot S_{j2} + \ldots + w_6 \cdot S_{j6} \quad (3)$$

Izraz (3) predstavlja zbroj produkata SHAP vrijednosti za određeni parametar *j*, dobivenih iz svakog od šest modela (S_{j1}, S_{j2}, ... S_{j6}) i njihovih pripadajućih težina (w_1, w_2, ... w_6), koje su određene proporcionalno točnosti vrijednosti R^2 svakog modela. Time se dobiva SHAP vrijednost koja integrira sve modele u jedinstvenu metriku važnosti.

Na slici 7. prikazane su dobivene vrijednosti utjecaja za svaki parametar, izražene pomoću ponderirane SHAP vrijednosti. Taj stupčasti grafikon omogućuje rangiranje parametara prema njihovoj ukupnoj važnosti kroz sve modele. Najviše vrijednosti zabilježene su za PGDS, LIMIT i K.Int. P.T., što upućuje na njihov dosljedan i dominantan utjecaj na predviđanja u svim modelima. Suprotno tome parametri poput KON.NAK., Max_Temp i H sign imali su najniže ponderirane SHAP vrijednosti, što upućuje na to da je njihov utjecaj na izlaznu varijablu bio minimalan ili ograničen na nekoliko modela. Taj dijagram omogućuje vizualnu procjenu ključnih čimbenika za buduće analize i potencijalno smanjenje broja varijabli.

Radi daljnje analize parametri su rangirani prema ponderiranim SHAP vrijednostima kako bi se postupno smanjio broj uključenih varijabli. Primjenom tog pristupa moguće je odrediti najutjecajnije parametre bez oslanjanja na tradicionalnu metodu postupne eliminacije, zahvaljujući integriranoj evaluaciji svih modela.

5.3. Optimiranje broja parametara i odabir najpreciznijeg modela

U toj fazi šest modela procijenjeno je s obzirom na njihovu sposobnost predviđanja indeksa Wi, s posebnim težištem na utjecaju smanjenja broja ulaznih parametara na točnost predviđanja (tj. R^2). Analiza je bila temeljena na podjeli podataka u omjeru 80/20, s parametrima koji su sekvencijalno uklonjeni prema njihovim unaprijed definiranim rangovima važnosti [41].

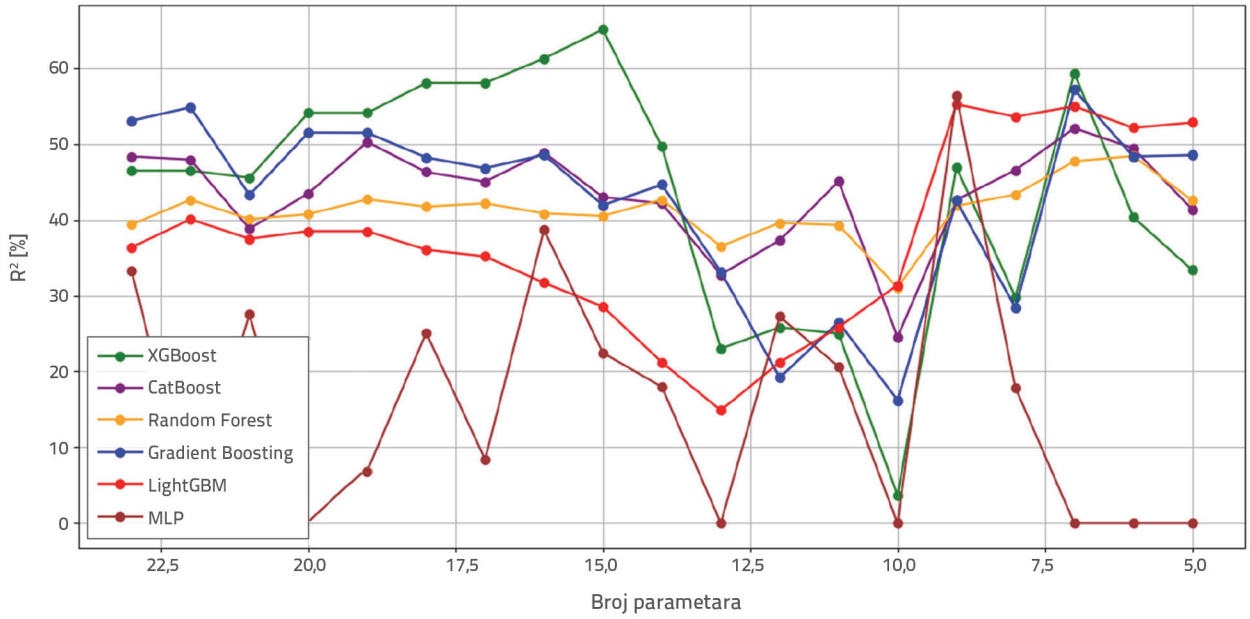
Za svaku iteraciju odabire se podskup najrelevantnijih značajki, nakon čega slijedi treniranje i testiranje modela za istu podjelu. MLP model uključivao je standardizaciju ulaza unutar *pipelinea*, dok je LightGBM model upotrebljavao posebno podešene hiperparametre za kontrolu složenosti modela. Upotrijebljen je fiksni *random_state* postavljen na 42 kako bi se osigurala ponovljivost rezultata.

Na slici 8. prikazana je varijacija vrijednosti R^2 za različit broj parametara za svaki model. Grafikon omogućuje vizualnu usporedbu osjetljivosti i robusnosti modela na smanjenje dimenzionalnosti. Važno je istaknuti da neki modeli poput GB-a održavaju stabilnu razinu učinkovitosti kroz širi raspon parametara, dok drugi poput MLP-a pokazuju nagle varijacije, osobito pri smanjenome broju ulaza.

U tablici 4. prikazana je maksimalna R^2 vrijednost postignuta za svaki model, zajedno s odgovarajućim brojem parametara, poredanima silaznim slijedom prema točnosti.

Tablica 4. Najveća vrijednost R^2 (%) i odgovarajući broj parametara za svaki model

Model	R^2 [%]	Broj parametara
XGBoost	65,05	15
GB	57,13	7
MLP	56,39	9
LightGBM	55,21	9
CatBoost	52,06	7
RF	48,40	6



Slika 8. Varijacija R^2 (%) s brojem ulaznih parametara za sve modele

Rezultati su pokazali da je XGBoost najučinkovitiji model za daljnju upotrebu, s najvišom postignutom vrijednošću R^2. Unatoč manjemu broju parametara, GB i CatBoost osiguravaju konkurentnu preciznost i stabilnost, što ih čini vrlo učinkovitima kada je količina podataka ograničena. MLP i LightGBM ostvarili su usporedive vrijednosti R^2, no samo pod određenim parametrima i s manjom dosljednošću kroz cijeli raspon.

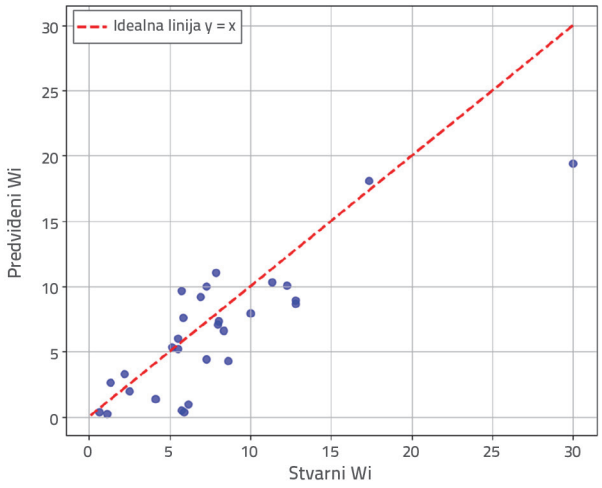
Ti rezultati upućuju na to da odabir odgovarajućeg modela i broja parametara može znatno poboljšati točnost predviđanja, čak i ako se ne oslanja na cijeli skup ulaznih varijabli. Ta analiza dodatno omogućuje usklađivanje dimenzionalnosti i stabilnosti modela, što je ključno za primjenu u praksi.

5.4. Testiranje i validacija

Proces testiranja i validacije ključan je za procjenu stabilnosti, točnosti i primjenjivosti razvijenog prediktivnog modela za Wi u stvarnim uvjetima. Kao što je to detaljno opisano u ovome odjeljku, za kvantificiranje točnosti predviđanja primijenjeno je testiranje temeljeno na regresiji, dok je validacija temeljena na klasifikaciji primijenjena za ispitivanje sposobnosti modela da identificira i rangira dionice s većim rizikom kako bi se podržale praktične odluke o intervencijama sigurnosti na cestama.

5.4.1. Testiranje (regresija, 80/20)

Za evaluaciju modela XGBoost, treniranog na 15 najutjecajnijih parametara prema SHAP vrijednostima, primijenjena je podjela podataka od 80 % za treniranje i 20 % za testiranje. Usporedba "predviđenog i stvarnog" s idealnom linijom $y = x$ omogućuje izravnu vizualnu procjenu podudarnosti između rezultata modela i promatranih vrijednosti Wi (slika 9.).



Slika 9. Usporedba predviđenih i stvarnih vrijednosti za Wi pomoću XGBoosta s podjelom podataka od 80 % za treniranje i 20 % za testiranje

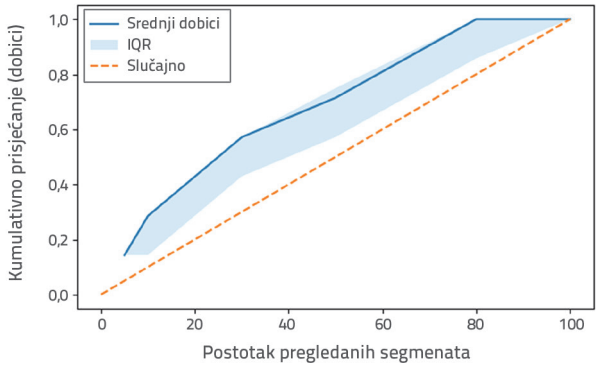
Većina točaka nalazila se u neposrednoj blizini idealne linije, uz očekivano, ali ograničeno raspršenje u ekstremnim vrijednostima. Kvantitativno, R^2 = 0,6505 upućuje na to da se znatan dio varijance Wi može pripisati modelu, dok vrijednosti MAE =2,72 i RMSE = 3,63 potvrđuju umjerene vrijednosti apsolutne pogreške i korijena srednje kvadratne pogreške. Zato ti rezultati podupiru primjenu modela XGBoost kao pouzdane osnove za operativnu procjenu rizika na razini dionica.

5.4.2. Validacija (klasifikacija za određivanje prioriteta rizika)

Osim preciznih regresijskih predviđanja za praktičnu primjenu ključno je da model da prioritet dionicama s najvećim rizikom na vrhu liste prioriteta. Regresijski model prilagođen je klasifikacijskoj postavci s pragom Wi ≥ 10,130 (prevalencija ≈ 20,5 %) kako bi se procijenio taj aspekt, a učinkovitosti su procijenjene kroz 100 *bootstrap* iteracija. Primjena metrika prilagođenih neuravnoteženim klasama (preciznost/odziv pri fiksnim stopama pregleda, PR-AUC) i dijagnostika usmjerenih na rangiranje (dobici/lift) metodološki je primjeren za zadatke određivanja prioriteta u analitici sigurnosti cestovnog prometa i usklađen je s nedavnim primjenama strojnog učenja u modeliranju rizika od sudara [42, 43]. Agregirani rezultati (medijani s 95 % CI) sažeti su kako slijedi:

- AUROC: 0,692 [0,519, 0,834]
- PR-AUC: 0,420 [0,229, 0,696]
- preciznosti pri 10 %: 0,500 [0,250, 0,750]
- odziv pri 10 %: 0,286 [0,143, 0,429]
- povećanje učinkovitosti (lift) pri 10 %: 2,857 [1,429, 4,286]

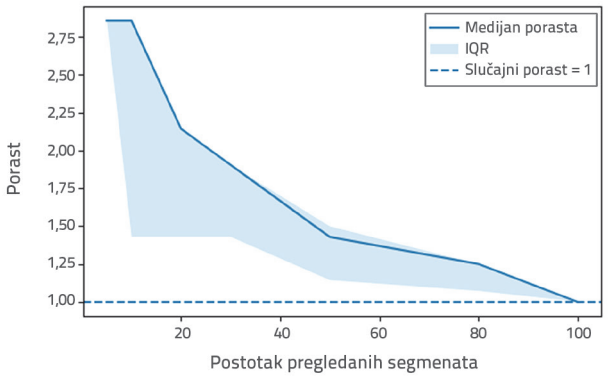
Na slici 10. prikazana je krivulja učinka (*gains*). Pregledom samo prvih 10 % cestovnih dionica primjenom ocjena rizika modela identificirana je otprilike polovina svih stvarno visokorizičnih dionica. To označava znatno poboljšanje u odnosu na slučajan odabir i jasan pokazatelj operativne korisnosti za rangiranje.



Slika 10. Krivulja učinka za prepoznavanje dionica s vrijednošću Wi ≥ 10, 130 (*bootstrap* medijan, 100 iteracija)

Na slici 11. prikazana je krivulja povećanja koja kvantificira prednost u odnosu na slučajan izbor. U prvih 10 % rangiranih dionica lift iznosi približno 2,9, što upućuje na to da model koncentrira znatno veći udio "pozitivnih" slučajeva pri vrhu liste, što je upravo ono što se očekuje od učinkovite prioritizacije.

Sveukupno, validacija temeljena na klasifikaciji pokazuje da model daje točna predviđanja za Wi te učinkovito rangira dionice prema riziku. U kombinaciji s regresijskim testovima ti nalazi pružaju uvjerljivu potvrdu da je model primjenjiv u sustavnome planiranju sigurnosti prometa.



Slika 11. Krivulja povećanja za procjenu relativnog učinka u rangiranome odabiru (*bootstrap* medijan, 100 iteracija)

6. Rasprava o rezultatima

6.1. Analiza rezultata

Analiza je obuhvatila šest modela (XGBoost, CatBoost, GB, RF, LightGBM i MLP) i provedena je u sklopu dva pristupa. Evaluacija temeljena na vrijednosti R^2 (R^2, MAE i RMSE) provedena je primjenom podjele 80/20 na treniranje i testni skup, uz objašnjivost modela SHAP-om i analizom važnosti permutacijom. Utjecaj varijabli izračunan je za svih šest modela i ponderiran prosjekom s težinama proporcionalnima R^2 vrijednosti svakog modela prema evaluacijskome protokolu podjele 80/20 (GB: R^2 = 0,5381, težina = 0,225; RF: R^2 = 0,5126, težina = 0,214; CatBoost: R^2 = 0,4664, težina = 0,195; LightGBM: R^2 = 0,3447, težina = 0,144; MLP: R^2 = 0,1607, težina = 0,067; i XGB: R^2 = 0,1043, težina = 0,044). Ta ponderirana agregacija SHAP vrijednosti stvara jedinstvenu rang-listu koja integrira sve modele, pri čemu je stabilna i manje osjetljiva na specifičnosti pojedinog algoritma.

Dobivene rang-liste pokazuju da su PGDS (AADT), LIMIT i K.Int. P.T. bili najdosljednije utjecajni čimbenici, pri čemu su PCI i Ave_Inc među vodećim infrastrukturnim/geometrijskim varijablama. Za razliku od njih, KON.NAK., H sign i Max_Temp pokazuju selektivnu ili nisku važnost. Mehanistički gledano, izloženost i radni uvjeti povećavaju osnovni rizik, dok ga uvjeti na kolniku i geometrija moduliraju trenjem, stabilnosti i vidljivosti.

Analiza ablacijske regresije (R^2 kao funkcija broja ulaznih varijabli) ukazala je na očitu ravnotežu između kompaktnosti modela i točnosti predviđanja. Vršne vrijednosti R^2 (%) i optimalan broj ulaznih varijabli po modelu jesu sljedeće: 65,05 (15 varijabli) za XGBoost, 57,13 (7 varijabli) za GB, 56,39 (9 varijabli) za MLP, 55,21 (9 varijabli) za LightGBM, 52,06 (7 varijabli) za CatBoost i 48,40 (6 varijabli) za RF. Time je dokazano da se najbolja generalizacija postiže primjenom reducirane, ali informativne podskupine varijabli, za razliku od primjene cijelog skupa ulaznih podataka.

Pri klasičnoj 80/20 podjeli podataka za validaciju XGBoost [47] s 15 ulaznih varijabli, odabranih kombinacijom SHAP i permutacijske važnosti, postigao je R^2 = 0,6505, MAE = 2,72

i RMSE = 3,63, pri čemu su testne točke bile koncentrirane oko idealne linije $y = x$, što potvrđuje snažno slaganje između predviđenih i opažanih vrijednosti Wi.

Za operativnu validaciju (prioritizaciju) izlaz regresijskog modela prilagođen je klasifikacijskome scenariju s pragom Wi ≥ 10,130 (prevalencija ≈ 20,5 %) te sažet preko 100x *bootstrap* iteracija. Rezultati su bili sljedeći: AUROC = 0,692, PR-AUC = 0,420, preciznost pri 10 % = 0,500, odziv pri 10 % = 0,286, i povećanje učinkovitosti pri 10 % = 2,857. Ti rezultati upućuju na to da je pregledom samo gornjih 10 % dionica obuhvaćen znatno veći udio stvarno visokorizičnih segmenata u usporedbi s nasumičnim odabirom (što je vidljivo i iz krivulja dobitka/povećanje učinkovitosti).

Ukratko, 15 ulaznih varijabli dovoljno je za stabilnu generalizaciju u modelu XGBoost, dok GB i CatBoost ostvaruju konkurentne performanse čak i s manjim brojem ulaza, što je posebno korisno u uvjetima ograničenih podataka. Osim što model predviđa Wi s visokim slaganjem izvan skupa za učenje (R^2), pristup učinkovito stavlja segmente s najvećim rizikom na vrh liste za inspekciju i intervenciju. Tumačenje se temelji na validiranim rezultatima izvan skupa za učenje. Rezultati dobiveni na cjelokupnome razvojnome skupu primijenjeni su isključivo u istraživačke svrhe, a ne za formalnu evaluaciju.

6.2. Usporedba s prethodnim istraživanjima

Validirani rezultati predikcije Wi prema podjeli 80/20 (XGBoost, 15 ulaznih varijabli: R^2 = 0,6505, MAE = 2,72, RMSE = 3,63) usklađeni su s aktualnim praksama koje preferiraju ansamble temeljene na stablima i ističu objašnjivost modela putem SHAP-a. U regresijskim istraživanjima sažetima u tablici 1. prijavljene vrijednosti R^2 obično se kreću oko 0, 84 do 0,92 za nacionalne ili urbane kontekste s većim i bogatijim skupovima podataka [15, 17, 18, 21, 23]. U hibridnim ili interpretacijski usmjerenim okružjima s težištem na urbanim sredinama rezultati se uglavnom kreću oko R^2 ≈ 0,83 do 0,87 [20]. Te su razlike očekivane i rezultat su razlike u ciljevima (dok mnoga istraživanja predviđaju ozbiljnost nesreća, ovo istraživanje modelira kontinuirani Wi), prostorne razlučivosti i opsega atributa. Zato su usporedbe ponajprije metodološke prirode, s težištem na simultanoj primjeni *boostinga* i objašnjive umjetne inteligencije (XAI), a ne na izravnoj brojčanoj usporedivosti.

U sklopu klasifikacijskih i hibridnih pristupa prethodna istraživanja obično su izvještavala o metrikama AUC, PR-AUC, preciznosti i odzivu, pri čemu se AUC kretao otprilike između 0,78 i 0,87, ovisno o specifičnome zadatku i skupu podataka [19, 20]. Radi operativne usporedivosti u ovom je istraživanju provedena i provjera prioritizacije: pri pragu Wi ≥ 10,130 (prevalencija ≈ 20,5 %; 100 × *bootstrap* iteracija) dobivene su sljedeće metrike: AUROC = 0,692, PR-AUC = 0,420, preciznost pri 10 % = 0,500, odziv pri 10 % = 0,286 i lift pri 10 % = 2,857, što upućuje na učinkovitu koncentraciju najrizičnijih dionica pri vrhu liste. Više R^2 i AUC vrijednosti u literaturi dijelom su posljedica većih i raznovrsnijih skupova podataka (prostorno i vremenski) koji omogućuju veću varijabilnost i učinkovitije učenje. U postavci sa 161 dionicom, uvjeti za generalizaciju prirodno su stroži.

Što se tiče odrednica, rezultati upućuju da izloženost i operativni uvjeti (PGDS/AADT, ograničenje brzine i gustoća prometa na raskrižju) imaju dominantan utjecaj, dok uvjeti kolnika i geometrija (PCI, uzdužni nagib) moduliraju rizik. Ti su rezultati u skladu sa istraživanjima koja kombiniraju *boosting* metode sa SHAP-om radi objašnjivosti [17, 20, 24]. Oni podupiru pristup ponderiranog kombiniranja SHAP vrijednosti preko više modela, istodobno opravdavajući smanjenje broja ulaznih varijabli bez znatnog gubitka sposobnosti generalizacije modela.

Ukratko, rezultati izvan skupa za učenje usklađeni su s aktualnim pristupima (ansambli modela u kombinaciji sa SHAP-om) te imaju operativnu primjenjivost za prioritizaciju. Tablica 1. služi kao referentni okvir za metodološki dosljednu usporedbu zadataka, metrika i skala.

7. Ograničenja i smjerovi budućih istraživanja

Iako su napredni modeli strojnog učenja postigli pouzdanu točnost pri predviđanju Wi izvan skupa za učenje (R^2 = 0,6505; MAE = 2,72; RMSE = 3,63), postoje određena ograničenja koja treba razmotriti.

Pozorno upravljanje rizikom od prekomjernog prilagođavanja neophodno je, posebno za modele koji uključuju velik broj parametara [48]. U validaciji temeljenoj na klasifikaciji koja se primjenjivala za određivanje prioriteta učinkovitost je bila ograničena prevalencijom klase (~20,5 %), što treba uzeti u obzir pri tumačenju vrijednosti AUROC/PR-AUC.

Glavno ograničenje proizlazi iz dostupnosti i detalja ulaznih podataka. Nedostaju ažurirane informacije o stanju vertikalnog i horizontalnog znakovlja za ceste, trenutnome stanju kolnika i potpunim klimatskim parametrima. Nadalje, podaci o prometnim nesrećama bili su ograničeni u smislu detaljnih opisa uzroka, uvjeta u vrijeme nesreća i točnih geografskih lokacija incidenata, što je utjecalo na preciznost modela.

Modeli su trenirani primjenom podataka iz glavne cestovne mreže. Zato bi primjena istog pristupa na druge kategorije cesta zahtijevala dodatnu prilagodbu.

Za buduća istraživanja preporučuje se proširenje baze podataka kako bi se uključile informacije o trenutnome stanju infrastrukture, specifičnijim klimatskim uvjetima i čimbenicima povezanima s ponašanjem sudionika u prometu. Osim toga primjena kombiniranih algoritama i objašnjivih tehnika strojnog učenja može dodatno poboljšati prediktivnu točnost i interpretabilnost rezultata modela.

8. Zaključak

Rezultati ovog istraživanja upućuju da napredni modeli strojnog učenja omogućuju pouzdano predviđanje ponderirang indeksa nesreća (Wi) na razini dionica i pružaju potporu u operativnome odlučivanju. U sklopu validacije prema podjeli 80/20 model XGBoost s 15 ulaznih varijabli odabranih permutacijskom

analizom SHAP postigao je vrijednosti R^2 od 0,6505, MAE od 2,72 i RMSE od 3,63, što upućuje na snažnu usklađenost predviđenih i promatranih vrijednosti izvan skupa za učenje. Kombinirana, ponderirana agregacija SHAP vrijednosti kroz sve modele omogućila je stabilno rangiranje determinanti, pri čemu su AADT (PGDS), ograničenje brzine (LIMIT) i gustoća raskrižja (K.Int. P.T) prepoznati kao najutjecajniji čimbenici, a slijede ih uvjeti kolnika (PCI) i uzdužni nagib (Ave_Inc.). Operativna provjera usmjerena na prioritizaciju dodatno je pokazala da, pri pragu Wi ≥ 10,130 (prevalencija ≈ 20,5 %), model učinkovito koncentrira rizik (AUROC = 0,692; PR-AUC = 0,420; preciznost pri 10 % = 0,500; odziv pri 10 % = 0,286; povećanje učinkovitosti pri 10 % = 2,857), što ga čini prikladnim za ciljane inspekcije i intervencije u uvjetima ograničenih resursa.

Glavni doprinosi ovog istraživanja su sljedeći: standardizirana usporedba modela ansambla i osnovnih modela izvan skupa za učenje, treniranih u jednakim uvjetima za predviđanje ponderiranog indeksa nesreća Wi; integrirani postupak rangiranja varijabli (ponderirani SHAP kroz šest modela) koji omogućuje smanjenje dimenzionalnosti bez znatnoga gubitka sposobnosti generalizacije te operativni okvir validacije koji povezuje rezultate regresije s praktičnom, rangiranom selekcijom visokorizičnih dionica.

Iako su rezultati obećavajući, oni i dalje ovise o granularnosti i pokrivenosti podataka (npr. detaljno stanje znakovlja na cesti, trenutno stanje kolnika, precizna geolokacija sudara i bogatiji klimatski deskriptori). Očekuje se da će proširenje i ažuriranje tih ulaznih podataka, uključivanje prostornih/vremenskih struktura i provođenje vanjske validacije na dodatnim mrežama poboljšati eksplanatornu vrijednost i prenosivost. U praksi rezultati upućuju na potrebu za kombiniranom strategijom upravljanja izloženošću i okruženjem brzine dopuštene u operativnome okolišu (npr. politike ograničenja brzine i upravljanje raskrižjima), uz održavanje i nadogradnju infrastrukture (stanje površine, odvodnja i vidljivost). Predložena metodologija pruža jasan, objašnjiv i praktično primjenjiv okvir za prioritizaciju sigurnosnih mjera na cestama u širokome opsegu.

LITERATURA

[1] World Health Organization: Road traffic injuries, Available at: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries (accessed March 2025)., 2024

[2] European Commission: Road Safety Statistics 2023, Available at: https://road-safety.transport.ec.europa.eu/european-road-safety-observatory_en (accessed March 2025).

[3] State Statistical Office of the Republic of North Macedonia: MakStat - Statistical Database. Available at: https://makstat.stat.gov.mk/PXWeb/pxweb/en/ (accessed March 2025), 2025

[4] Ziakopoulos, A., Yannis, G.: A review of spatial approaches in road safety, Accident Analysis & Prevention, 135 (2020), Paper 105323, https://doi.org/10.1016/j.aap.2019.105323

[5] Aguero-Valverde, J., Jovanis, P.P.: Spatial analysis of fatal and injury crashes in Pennsylvania, Accident Analysis & Prevention, 38 (2006) 3, pp. 618–625, https://doi.org/10.1016/j.aap.2005.12.006

[6] Silva, P.B., Andrade, M., Ferreira, S.: Machine learning applied to road safety modeling: A systematic literature review, Journal of Traffic and Transportation Engineering (English Edition), 7 (2020) 6, pp. 775–790, https://doi.org/10.1016/j.jtte.2020.07.004

[7] Almahdi, A., Al Mamlook, R.E., Bandara, N., Almuflih, A.S., Nasayreh, A., Gharaibeh, H., Alasim, F., Aljohani, A., Jamal, A.: Boosting ensemble learning for freeway crash classification under varying traffic conditions: A hyperparameter optimization approach, Sustainability, 15 (2023) 22, Paper 15896, https://doi.org/10.3390/su152215896

[8] Dong, S., Khattak, A., Ullah, I., Zhou, J., Hussain, A.: Predicting and analyzing road traffic injury severity using boosting-based ensemble learning models with SHAPley Additive exPlanations, International Journal of Environmental Research and Public Health, 19 (2022) 5, Paper 2925, https://doi.org/10.3390/ijerph19052925

[9] Kovačević, M., Ivanišević, N., Petronijević, P., Despotović, V.: Construction cost estimation of reinforced and prestressed concrete bridges using machine learning. Građevinar, 73 (2021) 1, pp. 1–13, https://doi.org/10.14256/JCE.2738.2019

[10] Gatarić, D., Ruškić, N., Aleksić, B., Đurić, T., Pezo, L., Lončar, B., Pezo, M.: Predicting road traffic accidents-Artificial neural network approach. Algorithms, 16 (2023) 5, Paper 257, https://doi.org/10.3390/a16050257

[11] Xiao, Y., Duan, Z.: An explainable multi-task deep learning framework for crash severity prediction using multisource data,Scientific Reports, 15 (2025), Paper 39226, https://doi.org/10.1038/s41598-025-09226-1

[12] Behboudi, N., Moosavi, S., Ramnath, R.: Recent advances in traffic accident analysis and prediction: A comprehensive review of machine learning techniques, arXiv preprint, arXiv:2406.13968. https://doi.org/10.48550/arXiv.2406.13968, 2024

[13] Li, W., Luo, Z.: Research on traffic accident risk prediction method based on spatial and visual semantics. ISPRS International Journal of Geo-Information, 12 (2023) 12, Paper 496, https://doi.org/10.3390/ijgi12120496

[14] Li, H., Chen, X.: Traffic accident risk prediction based on deep learning and spatiotemporal features of vehicle trajectories, PLOS ONE, 20 (2025) 7, Paper e0320656, https://doi.org/10.1371/journal.pone.0320656

[15] Chen, F., Liu, X.Q., Yang, J.J., Liu, X.K., Ma, J.H., Chen, J., Xiao, H.Y.: Traffic accident severity prediction based on an enhanced MSCPO-XGBoost hybrid model. Scientific Reports, 15 (2025), Paper 25729, https://doi.org/10.1038/s41598-025-00797-7

[16] Iranmanesh, M., Seyedabrishami, S., Moridpour, S.: Identifying high crash risk segments in rural roads using ensemble decision tree-based models. Scientific Reports, 12 (2022), Article 20024, https://doi.org/10.1038/s41598-022-24476-z

[17] Lee, J., Kim, S., Heo, T.Y., Lee, D.: Identifying the roadway infrastructure factors affecting road accidents using interpretable machine learning and data augmentation, Applied Sciences, 15 (2025, 5), Paper 501, https://doi.org/10.3390/app15020501

[18] Mengistu, A.K., Gedefaw, A.E., Baykemagn, N.D., Walle, A.D., Yehuala, T.Z., Alemayehu, M.A., Messelu, M.A., Assaye, B. T.: Predicting car accident severity in Northwest Ethiopia: A machine learning approach leveraging driver, environmental, and road conditions, Scientific Reports, 15 (2025), Paper 21913. https://doi.org/10.1038/s41598-025-08005-2

[19] Alshehri, A.H., Alanazi, F., Yosri, A.M., Yasir, M.: Comparing fatal crash risk factors by age and crash type using machine learning techniques, PLOS ONE, 19 (2024) 5, e0302171, https://doi.org/10.1371/journal.pone.0302171

[20] Ahmed, S., Hossain, M.A., Ray, S.K., Bhuiyan, M.M.I., Sabuj, S.R.: A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance, Transportation Research Interdisciplinary Perspectives, 19 (2023), Paper 100814, https://doi.org/10.1016/j.trip.2023.100814

[21] Megnidio-Tchoukouegno, M., Adedeji, J.A.: Machine learning for road traffic accident improvement and environmental resource management in the transportation sector (using UK STATS19 data), Sustainability, 15 (2023) 3, Paper 2014, https://doi.org/10.3390/su15032014

[22] Alpalhão, N., Sarmento, P., Jardim, B., de Castro Neto, M.: Assessing the risk of traffic accidents in Lisbon using a gradient boosting algorithm with a hybrid classification/regression approach, Transportation Research Interdisciplinary Perspectives, 21 (2025), Paper 101495, https://doi.org/10.1016/j.trip.2025.101495

[23] Guido, G., Shaffiee Haghshenas, S., Vitale, A., Astarita, V., Park, Y., Geem, Z.W.: Evaluation of contributing factors affecting number of vehicles involved in crashes using machine learning techniques in rural roads of Cosenza, Italy. Safety, 8 (2023) 2, Paper 28, https://doi.org/10.3390/safety8020028

[24] Xiao, Y., Duan, Z.: An explainable multi-task deep learning framework for crash severity prediction using multisource data, Scientific Reports, 15 (2025), Paper 9226. https://doi.org/10.1038/s41598-025-09226-1

[25] Public Enterprise for State Roads, Web-GIS Platform for Spatial Analysis and Visualization, Available at: http://62.77.137.99/pesr/webgis/#/map (accessed March 2025).

[26] Ministry of Local Self-Government, Development Program for Planning Regions for the Period 2021–2026, Government of the Republic of North Macedonia, 2021.

[27] Doncheva, R., Ognjenović, S.: Proektiranje patishta, University "Ss. Cyril and Methodius" – Faculty of Civil Engineering, Skopje, ISBN: 978-608-4510-60-4, 2024.

[28] Tobias, P., de León Izeppi, E., Flintsch, G., Katicha, S., McCarthy, R.: Pavement Friction for Road Safety: Primer on Friction Measurement and Management Methods, Federal Highway Administration (FHWA), Report No. FHWA-SA-23-007, 2023.

[29] Public Enterprise for State Roads, Web-GIS Platform for Spatial Analysis and Visualization, Available at: http://tdps.roads.org.mk/ (accessed March 2025).

[30] Gjeshovska, V., Taseski, G., Ilioski, B.: Intensive Precipitation in the Republic of North Macedonia, University "Ss. Cyril and Methodius" – Faculty of Civil Engineering, Skopje, ISBN: 978-608-4510-56-7, 2024.

[31] Government of the Republic of North Macedonia, Ministry of Transport, Project Implementation Unit, Handbook on Black Spot Management (BSM), July 2024.

[32] Murtagh, F., Contreras, P.: Algorithms for hierarchical clustering: an overview, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2 (2012) 1, pp. 86–97, https://doi.org/10.1002/widm.53.

[33] Trajkovski, V.: How to Select Appropriate Statistical Test in Scientific Articles, Journal of Special Education and Rehabilitation, 17 (2016) 3–4, pp. 5–28, https://doi.org/10.19057/jser.2016.7.

[34] Friedman, J., Hastie, T., Tibshirani, R.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics, https://doi.org/10.1007/978-0-387-21606-5, 2001.

[35] Kuhn, M., Johnson, K.: Applied Predictive Modeling, Springer, https://doi.org/10.1007/978-1-4614-6849-3, 2013.

[36] Raschka, S.: Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning, arXiv (2018), https://doi.org/10.48550/arXiv.1811.12808.

[37] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A.: CatBoost: unbiased boosting with categorical features, arXiv (2017), https://doi.org/10.48550/arXiv.1706.09516.

[38] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y.: LightGBM: A highly efficient gradient boosting decision tree, Advances in Neural Information Processing Systems 30 (NeurIPS 2017), pp. 3149–3157, URL: https://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree

[39] Kittler, J., Hatef, M., Duin, R.P.W., Matas, J.: On combining classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20 (1998) 3, pp. 226–239, https://doi.org/10.1109/34.667881

[40] Breiman, L.: Random Forests, Machine Learning, 45 (2001) 1, pp. 5–32,https://doi.org/10.1023/A:1010933404324.

[4'] Friedman, J.H.: Greedy Function Approximation: A Gradient Boosting Machine. Annals of Statistics, 29 (2001) 5, pp. 1189–1232, https://doi.org/10.1214/aos/1013203451.

[42] Ahmed, S., Hossain, M.A., Ray, S.K., Bhuiyan, M.M.I., Sabuj, S.R.: A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance, Transportation Research Interdisciplinary Perspectives, 19 (2023), Paper 100814, https://doi.org/10.1016/j.trip.2023.100814.

[43] Alshehri, A.H., Alanazi, F., Yosri, A.M., Yasir, M.: Comparing fatal crash risk factors by age and crash type using machine learning techniques, PLOS ONE, 19 (2024) 5, e0302171, https://doi.org/10.1371/journal.pone.0302171.

[44] Guyon, I., Elisseeff, A.: An Introduction to Variable and Feature Selection, Journal of Machine Learning Research, 3 (2003), pp. 1157–1182, https://jmlr.org/papers/v3/guyon03a.html.

[45] Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions, Advances in Neural Information Processing Systems, 30 (2017), pp. 4765–4774, https://doi.org/10.48550/arXiv.1705.07874

[46] Molnar, C.: Interpretable Machine Learning, Second edition, self-published, 2022, https://doi.org/10.1177/09726225241252009

[47] Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2016), https://doi.org/10.1145/2939672.2939785.

[48] Patil, P., Du, J.H., Kuchibhotla, A.K.: Bagging in overparameterized learning: Risk characterization and risk monotonization, arXiv preprint arXiv:2210.11445, (2022), https://doi.org/10.48550/arXiv.2210.11445.